CONCATENATIVE SPEECH SYNTHESIS USING A FINITE-STATE TRANSDUCER

Abstract

A method for concatenative speech synthesis includes a processing stage that selects segments based on their symbolic labeling in an efficient graph-based search, which uses a finite-state transducer formalism. This graph-based search uses a representation of concatenation constraints and costs that does not necessarily grow with the size of the source corpus thereby limiting the increase in computation required for the search as the size of the source corpus increases. In one application of this method, multiple alternative segment sequences are generated and a best segment sequence is then be selected using characteristics that depend on specific signal characteristics of the segments.